NVM Express Technical Errata

Errata ID	**033**
Change Date	**10/18/2012**
Affected Spec Ver.	**NVM Express 1.0d**
Corrected Spec Ver.	

Submission info

Name	Company	Date
John Carroll	Intel	9/20/2012
Peter Onufryk	IDT	9/20/2012
Eric Peterson	Synopsis	9/20/2012
Matthew Wilcox	Intel	9/20/2012
Judy Brock	Samsung	10/11/2012
Amber Huffman	Intel	10/11/2012
Santosh Singh	Samsung	10/12/2012
Ken Okin	Virident	10/18/2012
Kevin Marks	Dell	10/18/2012

Admin Command PRP field clarified for contiguous buffers larger than a page.

Invalid Queue Identifier when creating a Submission/Completion queue clarified.

Modified Cache Line Size PCIe register to align with PCIe spec.

Clarified Flush command to apply to only one namespace.

AER Capability changes.

Clarified Firmware Activate reset requirements

PCIe register changes to comply with PCIe 3.0

Fixed Compare & Write fused command typo

Description of the specification technical flaw:

Modify Figures 51 & 52 in section 5.9 as shown below:

Figure 51: Get Features – PRP Entry 1

Bit	Description
63:00	**PRP Entry 1 (PRP1):** ~~Specifies a data buffer that the Feature information shall be returned in if the Feature information is returned in a data structure. The buffer shall not have more than one physical discontinuity.~~ This field contains the first PRP entry, specifying the start of the data buffer. If no data structure is used as part of the specified feature, then this field is ignored.

Figure 52: Get Features – PRP Entry 2

Bit	Description
63:00	**PRP Entry 2 (PRP2):** This field contains the second PRP entry. ~~that specifies the location where data should be transferred to (if there is a physical discontinuity).~~ If PRP Entry 1 specifies enough space for the data structure, then this field is reserved. Otherwise, it specifies the remainder of the data buffer. This field shall not be a pointer to a PRP List as the data buffer may not cross more than one page boundary. If no data structure is used as part of the specified feature, then this field is ignored.

Modify Figures 55 & 56 in section 5.10 as shown below:

Figure 55: Get Log Page – PRP Entry 1

Bit	Description
63:00	**PRP Entry 1 (PRP1):** ~~Specifies a data buffer that the log page shall be returned to. The buffer shall not have more than one physical discontinuity.~~ This field contains the first PRP entry, specifying the start of the data buffer.

Figure 56: Get Log Page – PRP Entry 2

Bit	Description
63:00	**PRP Entry 2 (PRP2):** This field contains the second PRP entry. ~~that specifies the location where data should be transferred to (if there is a physical discontinuity).~~ If PRP Entry 1 specifies enough space for the data structure, then this field is reserved. Otherwise, it specifies the remainder of the data buffer. This field shall not be a pointer to a PRP List as the data buffer may not cross more than one page boundary.

Modify Figures 63 & 64 in section 5.11 as shown below:

Figure 63: Identify – PRP Entry 1

Bit	Description
63:00	**PRP Entry 1 (PRP1):** ~~Indicates a data buffer that the Identify data structure shall be returned to. The buffer shall not have more than one physical discontinuity.~~ This field contains the first PRP entry, specifying the start of the data buffer.

Figure 64: Identify – PRP Entry 2

Bit	Description
63:00	**PRP Entry 2 (PRP2):** This field contains the second PRP entry. ~~that specifies the location where data should be transferred to (if there is a physical discontinuity).~~ If PRP Entry 1 specifies enough space for the data structure, then this field is reserved. Otherwise, it specifies the remainder of the data buffer. This field shall not be a pointer to a PRP List as the data buffer may not cross more than one page boundary.

Modify Figures 70 & 71 in section 5.12 as shown below:

Figure 70: Set Features – PRP Entry 1

Bit	Description
63:00	**PRP Entry 1 (PRP1):** ~~Indicates a data buffer that the Identify data structure shall be returned to. The buffer shall not have more than one physical discontinuity.~~ This field contains the first PRP entry, specifying the start of the data buffer. If no data structure is used as part of the specified feature, then this field is not used.

Figure 71: Set Features – PRP Entry 2

Bit	Description
63:00	**PRP Entry 2 (PRP2):** This field contains the second PRP entry. ~~that specifies the location where data should be transferred to (if there is a physical discontinuity).~~ If PRP Entry 1 specifies enough space for the data structure, then this field is reserved. Otherwise, it specifies the remainder of the data buffer. This field shall not be a pointer to a PRP List as the data buffer may not cross more than one page boundary. If no data structure is used as part of the specified feature, then this field is not used.

Modify Figures 91 & 92 in section 5.14 as shown below:

Figure 91: Security Receive – PRP Entry 1

Bit	Description
63:00	**PRP Entry 1 (PRP1):** ~~Specifies a data buffer that contains the security protocol information. The buffer shall not have more than one physical discontinuity.~~ This field contains the first PRP entry, specifying the start of the data buffer.

Figure 92: Security Receive – PRP Entry 2

Bit	Description
63:00	**PRP Entry 2 (PRP2):** This field contains the second PRP entry. ~~that specifies the location where data should be transferred to (if there is a physical discontinuity).~~ If PRP Entry 1 specifies enough space for the data structure, then this field is reserved. Otherwise, it specifies the remainder of the data buffer. This field shall not be a pointer to a PRP List as the data buffer may not cross more than one page boundary.

Modify Figures 95 & 96 in section 5.15 as shown below:

Figure 95: Security Send – PRP Entry 1

Bit	Description
63:00	**PRP Entry 1 (PRP1):** ~~Specifies a data buffer that contains the security protocol information. The buffer shall not have more than one physical discontinuity and shall be 4KB minimum in size.~~ This field contains the first PRP entry, specifying the start of the data buffer.

Figure 96: Security Send – PRP Entry 2

Bit	Description
63:00	**PRP Entry 2 (PRP2):** This field contains the second PRP entry. ~~that specifies the location where data should be transferred to (if there is a physical discontinuity).~~ If PRP Entry 1 specifies enough space for the data structure, then this field is reserved. Otherwise, it specifies the remainder of the data buffer. This field shall not be a pointer to a PRP List as the data buffer may not cross more than one page boundary.

Modify a portion of Figure 36 in section 5.3.1 as shown below:

Figure 36: Create I/O Completion Queue – Command Specific Status Values

Value	Description
1h	**Invalid Queue Identifier:** The creation of the I/O Completion Queue failed due to an invalid queue identifier specified as part of the command. An invalid queue identifier is one that is currently in use or one that is outside the range supported by the controller

Modify a portion of Figure 40 in section 5.4.1 as shown below:

Figure 40: Create I/O Submission Queue – Command Specific Status Values

Value	Description
1h	**Invalid Queue Identifier:** The creation of the I/O Submission Queue failed due an invalid queue identifier specified as part of the command. An invalid queue identifier is one that is currently in use or one that is outside the range supported by the controller

Modify section 2.1.6 as shown below:

2.1.6 Offset 0Ch: CLS – Cache Line Size

Bits	Type	Reset	Description
07:00	~~RO~~ RW	00h	**Cache Line Size (CLS):** ~~Not supported by NVM Express.~~ Cache Line Size register is set by the system firmware or operating system to the system cache size.

Modify section 6.7 as shown below:

6.7 Flush command

~~The Flush command is used by the host to indicate that any data in volatile storage should be flushed to non-volatile media.~~

The Flush command shall commit data and metadata associated with the specified namespace(s) to non-volatile media. The flush applies to all commands completed prior to the submission of the Flush command. The controller may also flush additional data and/or metadata from any namespace.

All command specific fields are reserved.

Modify a portion of Figure 66 in section 5.11 as shown below:

Bytes	O/M	Description
525	M	**Volatile Write Cache (VWC):** This field indicates attributes related to the presence of a volatile write cache in the implementation. Bits 7:1 are reserved. Bit 0 if set to '1' indicates that a volatile write cache is present. If cleared to '0', a volatile write cache is not present. If a volatile write cache is present, then the host may issue Flush commands and control whether it is enabled with Set Features specifying the Volatile Write Cache feature identifier. If a volatile write cache is not present, ~~the host shall not submit~~ Flush commands complete successfully and have no effect, and ~~nor~~ Set Features commands ~~or Get Features~~ with the Volatile Write Cache identifier field set shall fail with Invalid Field status.

Modify section 2.6.1 as shown below:

2.6.1 Offset AERCAP: AERID – AER Capability ID

Bits	Type	Reset	Description
31:20	RO	Impl Spec	**Next Pointer (NEXT):** Indicates the next item in the list. This may be other capability pointers or it may be the last item in the list.
19:16	RO	Impl Spec ~~2h~~	**Capability Version (CVER):** Indicates the version of the capability structure. Reset value may be 1h or 2h.
15:0	RO	0001h	**Capability ID (CID):** Indicates that this capability structure is an Advanced Error Reporting capability.

Modify a portion of section 5.7.1 as shown below:

5.7.1 Command Completion

A completion queue entry is posted to the Admin Completion Queue if the controller has completed the requested action (specified in the Activate Action field). For requests that specify activation of a new firmware image and return with status code value of 00h, any controller level reset defined in section 7.3.1 activates the specified firmware. Firmware Activate command specific status values are defined in Figure 46.

Modify a portion of section 7.2.5.2 as shown below:

- CMD1.CDW12.NLB is set to 3h ~~4h~~, indicating that ~~4h LBAs~~ four logical blocks of a size of 4KB each are to be compared against. This value shall be the same as CMD0.CDW12.NLB.

Modify a portion of section 2.3.2 as shown below:

2.3.2 Offset MSICAP + 2h: MC – Message Signaled Interrupt Message Control

Bits	Type	Reset	Description
15:09	RO	0	Reserved
08	RO	Impl Spec ~~0~~	**Per-Vector Masking Capable (PVM):** Specifies whether controller supports MSI per-vector masking.

Modify a portion of section 2.5.6 as shown below:

2.5.6 Offset PXCAP + Ch: PXLCAP – PCI Express Link Capabilities

Bits	Type	Reset	Description
31:24	RO	HwInit	**Port Number (PN):** This field ~~indicates~~ specifies the PCI Express port number for this device.
23~~:22~~	RO	0h	Reserved
22	RO	HwInit	**ASPM Optionality Compliance (AOC):** This field specifies Active State Power Management (ASPM) support

Modify a portion of section 3.1.5 as shown below:

Bit	Type	Reset	Description
00	RW	0	**Enable (EN):** When set to '1', then the controller shall process commands based on Submission Queue Tail doorbell writes. When cleared to '0', then the controller shall not process commands nor post completion queue entries to Completion Queues. When this field transitions from '1' to '0', the controller is reset (referred to as a Controller Reset). The reset deletes all I/O Submission Queues and I/O Completion Queues, resets the Admin Submission Queue and Completion Queue, and brings the hardware to an idle state. The reset does not affect PCI Express registers nor the Admin Queue registers (AQA, ASQ, or ACQ). All other controller registers defined in this section and internal controller state (e.g., Feature values defined in section 5.12.1 that are not persistent across power states) are reset to their default values. The controller shall ensure that there is no data loss for commands that have had corresponding completion queue entries posted to an I/O Completion Queue prior to the reset operation. Refer to section 7.3 for reset details. When this field is cleared to '0', the CSTS.RDY bit is cleared to '0' by the controller once the controller is ready to be re-enabled. When this field is set to '1', the controller sets CSTS.RDY to '1' when it is ready to process commands. Setting this field from a '0' to a '1' when CSTS.RDY is a '1,' or setting this field from a '1' to a '0' when CSTS.RDY is a '0,' has undefined results. The Admin Queue registers (AQA, ASQ, and ACQ) shall only be modified when EN is cleared to '0'.

Modify a portion of section 3.1.1 as shown below:

Bit	Type	Reset	Description
31:24	RO	Impl Spec	**Timeout (TO):** This is the worst case time that host software shall wait for the controller to become ready (the larger of the time required when CSTS.RDY is set to '1' or CSTS.RDY is cleared to '0') after a power-on or reset (section **Error! Reference source not found.**). This worst case time may be experienced after an unclean shutdown; typical times are expected to be much shorter. This field is in 500 millisecond units.

Modify a portion of section 7.6.1 as shown below:

7.6.1 Initialization

The host should perform the following actions in sequence to initialize the controller to begin executing commands:

1. Set the PCI and PCI Express registers described in section **Error! Reference source not found.** appropriately based on the system configuration. This includes configuration of power management features. Pin-based or single-message MSI interrupts should be used until the number of I/O Queues is determined.

2. The host waits for the controller to indicate that any previous reset is complete by waiting for CSTS.RDY to become '0.'

~~2~~3. The Admin Queue should be configured. The Admin Queue is configured by setting the Admin Queue Attributes (AQA), Admin Submission Queue Base Address (ASQ), and Admin Completion Queue Base Address (ACQ) to appropriate values.

Modify a portion of section 3.1.6 as shown below:

Bit	Type	Reset	Description
00	RO	0	**Ready (RDY):** This field is set to '1' when the controller is ready to accept Submission Queue Tail doorbell writes~~process commands~~ after CC.EN is set to '1'. This field shall be cleared to '0' when CC.EN is cleared to '0'. Commands shall not be submitted to the controller until this field is set to '1' after the CC.EN bit is set to '1'. Failure to follow this requirement produces undefined results. Host software shall wait a minimum of CAP.TO seconds for this field to be set to '1' after setting CC.EN to '1' from a previous value of '0'.

Disposition log

9/20/2012	Erratum captured.
10/3/2012	Expanded changes to PRP Entry to refer to each command individually. Added reference to NVM subsystem in flush. Removed applying this ECN to NVMe 1.1 (content will be copied to ECN 001). Added change to AER capability.
10/4/2012	Added clarification to flush command completion.
10/11/2012	Added Command completion clarification, and modified flush section
10/12/2012	Added clarification for toggling CC.EN, modified PXLCAP & MSICAP, reworded firmware activate and flush changes, fixed typo in write portion of compare & write
10/18/2012	Modified CC.EN change – removed modification to 3.1.6, added modification to 3.1.5
10/31/2012	Modified language for 3.1.5 and added clarification in 7.6.1 and CAP.TO
11/1/2012	Added command processing language
12/6/2012	Erratum ratified.